Face Recognition and Event Detection in Video: An Overview of PROVE-IT Projects

Prepared by:
David Bissessar
Ehren Choy
Dmitry Gorodnichy
Tony Mungham
Science & Engineering Directorate, Canadian Border Services Agency
79 Bentley Avenue, Ottawa, Ontario K1A 0L8 Canada

Scientific Authority:
Pierre Meunier
DRDC Centre for Security Science
613-944-4367

Contract Report
DRDC-RDDC-2014-C167
July 2014

IMPORTANT INFORMATIVE STATEMENTS

PROVE-IT (FRiV) Pilot and Research on Operational Video-based Evaluation of Infrastructure and Technology: Face Recognition in Video, PSTP 03-401BIOM was supported by the Canadian Safety and Security Program (CSSP) which is led by Defence Research and Development Canada's Centre for Security Science, in partnership with Public Safety Canada. Led by Canada Border Services Agency partners included : Royal Canadian Mounted Police, Defence Research Development Canada, Canadian Air Transport Security Authority, Transport Canada, Privy Council Office; US Federal Bureau of Investigation, National Institute of Standards and Technology, UK Home Office; University of Ottawa, Université Québec (ÉTS).

The CSSP is a federally-funded program to strengthen Canada's ability to anticipate, prevent/mitigate, prepare for, respond to, and recover from natural disasters, serious accidents, crime and terrorism through the convergence of science and technology with policy, operations and intelligence.

CBSA ASFC

PROTECTION • SERVICE • INTEGRITY   PROTECTION • SERVICE • INTÉGRITÉ

Science and Engineering Directorate

Applied Research and Development Division

Division Report : 2013-04
June, 2013

Face Recognition and Event Detection in Video: An Overview of PROVE IT Projects (BIOM401 and BTS402)

David Bissessar
Ehren Choy
Dmitry Gorodnichy
Tony Mungham

Canada Border Services Agency   Agence des services frontaliers du Canada

Canada

Table of Contents

Section 1. Introduction

Executive Summary

This report gives an overview of the BIOM401 and BTS402 projects which took place from June 2011 to May 2013. The report summarizes the achievements, key outputs, and the main findings of each project. The appendices of this document include summary tables which are extracted from the main project reports.

These projects measured the maturity levels of technologies enabling extraction of information from video footage, with BIOM401 focused on face recognition and BTS402 focused on event detection in video.

The projects resulted in an increased capability to make recommendations and investment decisions for deployment or further research into these technologies, taking into account the difficulties present in different types of operational environments and the limitations of key functionalities in each of these settings. As secondary outputs, the projects produced technology demonstrations, refereed publications, and an alternative assessment scale in support of the main project findings.

The outcomes of these projects should be meaningful for all who seek to effectively take action based on video information as events occur, or extract information and intelligence from the vast amounts of collected video footage

Overview of Projects

The BIOM401 and BTS402 projects took place from June 2011 to May 2013. These projects were funded by the Centre for Security Science (CSS) of Defence Research and Development Canada (DRDC), and were led by the Video Surveillance and Biometric (VSB) group of the Science and Engineering Directorate of the Canada Border Services Agency.

The projects measured the maturity levels of technologies for face recognition and event detection in video. The studies defined different operational environments (kiosk, interview counter, chokepoint, and large hall) and examined the readiness level of academic algorithms and commercial solutions in each of those environments.

BIOM401 assessed the readiness of a number of face recognition functionalities (such as face detection, person tracking between cameras, fusion of biometric modalities) in indoor environments. BTS402 focussed on readiness of the event detection functionalities (such as baggage left behind, tail-gating, and camera tampering) in both indoor and outdoor environments. The study highlights performance expectations, and deployment timelines within each operational environment.

The project partners for BIOM401 included L’Ecole de technology supérieure (ETS) in Montreal, and the TAMALE research group from Ottawa University, specialising in biometrics and machine learning, respectively. The partners on BTS402 were The University of Ottawa`s VIVA lab, and the Centre de recherche en informatique a Montreal (CRIM) lab in Montreal, both specialists in video analytics.

Section 2. Outputs and Achievements

Both projects produced the primary output of increased ability to recommend technologies. In support of the primary output, the projects also produced a number of secondary outputs, structured to provide clear value as independent deliverables. For both projects, the secondary outputs include a) readiness level assessments, b) an assessment framework, c) executable technology demonstrations d) scientific publications and e) strengthened inter-agency ties. This section summarizes both primary and secondary outputs.

Primary Outputs

Both projects have the common primary output of increased knowledge in their key areas, and increased ability to make recommendations.

1) **Increased ability to recommend face recognition solutions:** The primary output of project BIOM401 has been an increased capacity to recommend and critique proposed systems for Face Recognition in Video, for which the findings are discussed in Section 4 of this report.

2) **Increased ability to recommend event detection solutions:** As above, project BTS402 has improved the ability of the community of practice to propose feasible solutions and architectures for their appropriate operating environments. The main findings of BTS402 are presented in Section 5 of this report.

Secondary Outputs

A number of secondary outputs have been created by the projects which contribute to the increased ability to make recommendations, but also have high value in their own right as stand-alone deliverables.

3) **Assessment framework:** The projects present an assessment framework based on the Technology Readiness Level (TRL) scale defined by the United States Department of Defense (see http://www.acq.osd.mil/chieftechnologist/publications/docs/TRA2011.pdf). An alternative assessment framework was defined and used which expresses readiness in terms of required research and development, forecasted time to deployment, and required internal technical capability. The technologies have been assessed in environments of increasing difficulty.

4) **Technology products acquisition and verification:** The funding and mandate of the PROVE-IT projects have allowed the purchase and evaluation of commercial products such as Cognitec (for face recognition), and specialized Bosch cameras and encoders (for video analytics).

5) **Technology demonstrations:** Ten technology demonstrations were developed as part of the project to verify the limits of commercial and academic algorithms.

6) **Enhanced scientific capability:** The funding and mandate provided by the projects resulted in increased internal skills and knowledge which will continue to benefit the CBSA and the Government of Canada.

7) **Strengthened inter-agency ties:** The projects have promoted inter-agency ties within the Government of Canada, and the international community (UK HomeOffice, FBI, NIST).

Section 3. Assessment Framework

The projects assessed the technical capacity of commercial and academic algorithms using an assessment scale that expresses readiness in terms of timelines for deployment, and the level and type of technical effort required to deploy the technology. Required technical effort may include product configuration by Information Technology specialists, operationally-focused tuning and verification of commercial technologies, specialized algorithms developed by applied research and development and/or engineering groups, or research of a more exploratory nature that is more appropriately conducted by academia. This assessment framework describes four levels of readiness that are mapped to one or more of the nine levels within the Technology Readiness Level (TRL) scale[1].

1) **Operational Functionalities:** Functionalities in this category are the most mature and most straight forward. Often we have a commercial product with a documented install base. Depending on the environment type, functionalities in this category may be deployed within one year with little or no customization and predictable results. The technical effort for deployment can be achieved through vendor support, or with the help of IT solutions and engineering groups. This category of technologies represents TRL 7-9 maturity.

2) **Short-term Applied Research and Development Functionalities:** Functionalities in this category are reaching a high level of maturity. Technologies of this maturity level can be expected to be deployed in up to two years with a moderate investment in applied research and development. This category maps to levels 5-6 in the TRL scale.

3) **Medium-term Applied Research and Development Functionalities:** This category features functionalities that are challenging, but still deployable under the right circumstances in a 3-5 year timeframe. Technologies of this maturity level can be deployed with a significant applied research and development investment where the government agency works with industry and academia towards operationally meaningful functionalities. Technologies in this category are TRL level 4.

4) **Academic Research Functionalities:** Functionalities with a deployment timeline over five years tend to be the domain of academia and small technology start-ups. Government agencies should continue to observe this segment for breakthroughs, and perhaps influence direction. Technologies in this category are of TRL level 1-3.

The assessment scale provides context for findings discussed in Section 4 and Section 5, and assist when consulting the summary tables found in Appendix 1 and Appendix 2.

---

[1] The Technology Readiness Level (TRL) scale is defined by the United States Department of Defense (see http://www.acq.osd.mil/chieftechnologist/publications/docs/TRA2011.pdf)

Section 4. Key Findings: BIOM401 – Face Recognition in Video

Overview

The Face Recognition in Video project identified 14 functionalities involving aspects of face recognition in video and four types of operational environments where these functionalities could be deployed. The technological readiness of each of these functionalities was assessed against the four types of deployment environments. This section presents a summary of the results of this project. This section can be read in parallel to the summary table found in Appendix 1.

Environment Types

The environment types presented in the Face Recognition in Video project are defined as follows:

> A **Type 0 (kiosk) environment** features a cooperative traveller well positioned (essentially stationary) in close proximity to the capture camera, and willing to follow cues to assist the system in obtaining a high quality facial image. In this setting, we can expect images near passport-quality to be captured at verification time. Examples of Type 0 environments include **kiosks** and **e-Gates**.
>
> A **Type 1 (interview counter) environment** features close proximity of the camera to the subject, but less control over an individual's position. In a Type 1 environment, the subject is engaging in a stationary (yet naturally dynamic) activity such as an interview at a counter. The subject is unconcerned or unaware that a biometric sample is taken. Examples of Type 1 environments include **primary** and **secondary processing areas**.
>
> A **Type 2 (chokepoint) environment** features travellers on the move, walking at varying speeds through a building's "chokepoint" that channels traffic in a predictable manner without affecting its flow. The gait and pose angle of a subject cannot be controlled, but the path of the subject is predictable. Occlusion, due to crossing of paths, may occur but are considered as exceptions rather than normal occurrences. Examples of Type 2 environments include **hallways**, **doorways** and **turnstiles**.
>
> A **Type 3 (large hall) environment** is the most challenging environment for face recognition analyzed in this study. A Type 3 environment is an indoor environment where a number of subjects are freely moving about. There is no assumption of proximity or direction of motion, and occlusion is frequent. Examples of Type 3 environments include **waiting areas** and **baggage claim areas**.

Only indoor environments are examined in this report. For each functionality, it is assumed that a purpose-built camera is used with optimal placement and lighting. It is further assumed that a subject is acting in a natural manner within the scene and is not actively seeking to deceive the system.

It should be noted that the assessment of functionalities in the Type 0 environment was not within scope of this project because it typically requires the use of still imaging versus video imaging which is used in the other environments. While the Type 0 environment is out of scope, it is appreciated that the reader can derive meaningful information by comparing the Type 0 environment with the other three environments. For the purposes of this study, therefore, we have estimated the readiness of the target technologies in the Type 0 environment to be greater or equal to the readiness of the technology in the Type 1 environment. This is based on knowledge of the operational, technical, and commercial domains. However, it remains an estimate as no empirical studies have been conducted under the scope of this project. Further investigation into e-Gates is the subject of another project currently led by CBSA and funded under DRDC's Centre for Security Science Program.

Functionalities

The following 14 functionalities were assessed by the Face Recognition in Video project:

Detection Functionalities

1) **Face Detection in Live Video:** For this functionality, the system detects the presence of faces within a video, and gives the location of each faces. Face detection is a first step in the face recognition process. This functionality is deployment ready in Type 0, Type 1 and Type 2 environments, and medium-term applied research and development in a Type 3 environment.

2) **Face Extraction from Archive Video:** In this functionality, the system extracts face images from archived video. This functionality is analogous to functionality 1) above, "face detection in live video", except that historic footage is searched rather than a real-time video stream from a camera. This functionality is deployment ready in Type 0, Type 1 and Type 2 environments, and medium-term applied research and development in a Type 3 environment.

Tracking Functionalities

3) **Face Tracking Across a Single Video:** The system determines the path of a person within a video sequence. A face is detected within a starting frame, and the surrounding shaped is tracked through subsequent video frames. This level of functionality does not perform face recognition, but uses functionality 1), above. This functionality is deployment ready in Type 0, Type 1 and Type 2 environments and academic research in the Type 3 environment.

4) **Face Tracking Across Multiple Videos:** In this functionality, the system determines the path of a person passing through video streams which overlap on a surveillance area. Here some facial similarity can be used in this functionality to ensure confidence that the tracked individual is the same between video streams. This functionality is deployment ready in Type 0, Type 1 and Type 2 environments but is in the realm of academic research in the Type 3 environment.

Recognition

5) **Face Recognition for Watch List Screening (Binary):** Binary watch-list screening returns the match status of an individual in a video stream against a set of images of persons of interest. The system provides an alert if the person in the video is sufficiently similar to any of images in the set. The matched image is often returned. This functionality is short-term applied research and development in a Type 0 environment, medium term applied research and development in a Type 1 environment, and is in the realm of academic research in the Type 2 and Type 3 environments. [2]

6) **Face Recognition for Watch List Screening (Triaging):** Watch-list triaging extends the functionality in item 5), above adding levels of match confidence which are based on similarity of match, and image quality. This system is intended to aid a human operator in making an decision on person identification, and operational next steps. Confidence bands can be color-coded to quickly attract attention. This functionality is short-term applied research and development in a Type 0 environment, medium-term applied research and development in Type 1 and Type 2 environments, and is in the realm of academic research in the Type 3 environment.

[2] Watch-list screening is different from e-gate functionality. For e-Gates, a facial image captured is compared in a one-to-one manner to the biometric captured for the claimed identity at enrolment-time. In watch-list screening, a facial image captured at the time of passage is compared in a one-to-many manner, to a set of images for different subjects, and an alert is signalled if sufficient similarity is found.

7) **Face Recognition from Multiple Videos:** This functionality determines the presence of a person in multiple videos, taken from different cameras on differing scenes. Here, the camera angle and lighting can vary significantly from one camera to the other, and there is not necessarily a high quality reference image although the reference video stream is selected to be that with the best quality face images. This functionality is deployment ready in a Type 0 environment, short-term applied research and development in a Type 1 environment, medium-term applied research and development in a Type 2 environment, and is in the realm of academic research in the Type 3 environment.

8) **Face Fusion from Multiple Videos:** Given one or more video sequences of a single person, the system combines information from all video frames together to improve matching capability. This functionality is medium-term applied research and development in Type 0 and Type 1 environments, and is in the realm of academic research in the Type 2 and Type 3 environments.

Association

9) **Assisted Face Tagging and Grouping using Visual Analytics:** In this functionality, the system gathers similar face images from multiple videos and presents the results to a user for analysis. This functionality can be used on archival footage to retrieve video segments in which a person of interest appears for verification by a human operator. This functionality is short-term applied research and development in Type 0, Type 1 and Type 2 environments, and medium term applied research and development in a Type 3 environment.

10) **Automated Face Tagging and Grouping:** In this functionality, the system associates similar face images together across videos from multiple scenes. A human operator may validate and refine the associations provided by the system, and initiate system recalculation of associations based on new information. This functionality can be used to build intelligence on the presence of persons of interest across scenes. This functionality is short-term applied research and development in Type 0 and Type 1 environments, medium term applied research and development in a Type 2 environment, and is in the realm of academic research in the Type 3 environment.

Soft Biometrics

11) **Facial Expression Analysis:** The system determines the facial expression of persons in a video. This functionality is deployment-ready in a Type 0 environment, medium-term applied research and development in Type 1 and 2 environments, and is in the realm of academic research in the Type 3 environment.

12) **Human attribute recognition (Gender/Age/Race)**
The system determines the gender/age/race of a person in a video based on analysis of the face. This functionality is medium-term applied research and development in Type 0, 1 and 2 environments, and is in the realm of academic research in the Type 3 environment.

Multiple Biometrics

13) **Face Recognition to Improve Voice/Iris Biometrics**: Here, the system uses face recognition as a supplementary biometric to increase confidence on a match made using a different biometric (for example iris, voice, or fingerprints). This functionality is deployment-ready in a Type 0 environment, medium-term applied research and development in a Type 1 environment, and is in the realm of academic research in the Type 2 and Type 3 environments.

14) **Soft biometrics to improve face recognition**

The system uses a “soft” biometric (a biometric feature which does not uniquely identify such as inferred person height, or gait) as additional information to improve an identification decision based on facial similarity. This functionality is medium-term applied research and development in Type 0, 1 and 2 environments, and is in the realm of academic research in the Type 3 environment.

Overall, the success of using the technology depends on the degree of difficulty imposed by the deployment environment. In Type 0 environments, the study finds there are mature commercial products for face detection that are ready for immediate deployment. However, in Type 3 environments these functionalities become difficult research problems requiring a longer timeline for maturity.

Section 5. Key Findings: BTS402 – Event Detection in Video

Overview

The Event Detection in Video project identified seven types of behavioural events involving persons, baggage, or crowds and defined five types of operational environments where these technologies could be deployed. This section presents a summary of the results of this project. This section can be read in parallel to the summary table found in Appendix 2.

Environment Types

The Event Detection in Video project defines operational environments, in much the same manner as the Face Detection project, with the additional distinction between single person and multiple person chokepoint environments, as well as the addition of an outdoor environment.

For Event Detection, the environment types are defined as Type 1, Type 2, Type 3, Type 4 and Type 5, and are characterized as follows:

A **Type 1 (interview counter) environment** features close proximity of the camera to the subject, and a subject who is relatively stationary in the scene. The subject is unconcerned or unaware that a biometric sample is taken. Examples of Type 1 environments include **primary** and **secondary processing areas**.

A **Type 2 (chokepoint) environment** features travellers on the move, walking at varying speeds through a building “chokepoint” that channels traffic in a predictable manner without affecting its flow. The gait and pose of a subject cannot be controlled, but the path of the subject is predictable. Occlusion may occur, but are considered as exceptions rather than normal occurrences. Examples of Type 2 environments include **hallways**, **doorways** and **turnstiles**.

A **Type 3 (multi-person chokepoint) environment** is similar to a Type 2 environment, but may include many travellers walking at varying speeds going through the chokepoint. As above, the gait and pose of the subjects is not controlled, and the flow of people through the environment remains predictable. Occlusions occur more frequently here than in a Type 2 environment. Examples of Type 3 environments include **busy hallways** and **doorways**.

A **Type 4 (large hall) environment** is an indoor environment where a number of subjects are freely moving about. There is no assumption of proximity or direction of motion, and occlusion due to crossing of paths is frequent. Examples of Type 4 environments include **waiting areas** and **baggage claim areas**.

A **Type 5 (outdoor) environment** is an outdoor environment where a number of subjects are freely moving about. There is no assumption of proximity or direction of motion, and occlusion due to crossing of paths is frequent. Examples of Type 5 environments include **parking lots** and **sidewalks**.

Functionalities

The BTS402 project defines seven categories of functionality to be analyzed: 1) Unattended Baggage Detection; 2) Tailgaiting Detection; 3) Person Tracking; 4) Person-baggage Tagging; 5) Loitering Detection; 6) Crowd Tracking; and 7) Camera Tampering. A description of the project findings for each category of functionality follows:

1) **Unattended Baggage Detection:** Unattended baggage testing aimed to examine the scenario of possible suspicious items left in a common area. There are a number of variables involved in the scenario, including the size of the object left behind, and the amount of time after which it is determined to be abandoned. The BTS402 project determined that this type of functionality is largely in the domain of applied research and development. The study noted that possible deployment ready systems may exist depending on the degree to which an organisation is prepared to accept constraints with respect to the relative size of the object in the scene, the degree of motion in the scene, and the overall scene viewing conditions.

2) **Tailgaiting Detection:** This functionality attempts to detect events involving two entities where the second entity follows the first one closely through an access control point in an attempt to obtain unauthorized entry. The BTS402 project determined that in Type 1 environments, tailgating detection can achieve reasonable success rates with about 12 months of development investment. The functionality remains in the applied research and development domain in Type 2 environments. For more challenging environments, the results are varied and this functionality is in the realm of academic research.

3) **Person Tracking:** Person tracking applications attempt to follow a person as moving through a scene. Variations on this functionality include person counting, person tracking in the presence of running, and opposite direction tracking. In Type 1 environments, person tracking may be deployable with up to one year with moderate customization and configuration by Information Technology personnel. With small crowds (such as a multi person choke-point), this functionality becomes a harder task, requiring applied research and development. In an open crowd environment (such as Type 3 waiting area environments) this functionality is in the domain of academic research.

4) **Person-baggage Tagging**: Person-baggage tagging consists of maintaining the association between a person and their handheld baggage, and holding that association even when the baggage is deposited, and the person moves around the room. This functionality is medium-term applied research and development in Type 1 environments and Academic Research in Type 2, 3, 4, and 5.environments.

5) **Loitering Detection:** Loitering detection refers to the use of video analytics to determine if a person stays in the same area for a certain period of time. This functionality can be ready for operational deployment in Type 1 environments in up to one year. If the loitering area can be confined to a localized region of interest (an “improper standing area”) then this functionality is short-term applied research and development investment in Type 2,3, and 4 environments. If the allowable area a subject may wander becomes less constrained, the functionality becomes the subject of academic research in Type 2, 3, and 4 environments.

6) **Crowd Tracking:** Video analytics can be used in to determine patterns in the movement of groups of people. This includes crowd splitting, crowd merging, crowd density estimation, crowd formation and rapid dispersion. This functionality is most applicable in open areas (and to a lesser extent multi-person chokepoints). Specific tasks (such as detecting crowd splitting or merging) are deployment-ready with customizations in a Type 1 environment. Other tasks can expect results with a short-term investment for applied research and development.

7) **Camera Tampering:** Camera tampering uses video analytics to determine camera reliability by monitoring changes in the scene it views. Tampering events can include disconnected cameras or obstructed cameras. Technology to detect tampering of camera is mature across all deployment environments considered by the study.

Appendix 1: Face Recognition in Video Evaluation Matrix

Face Recognition in Video Environment Types

TYPE	DEFINITION
0	Cooperative biometric setup such as in Access Control or e-Gate. (Outside the scope of this project).
1	Semi-constrained setup, such as in Primary Inspection Lanes (PIL).
2	Unconstrained, free-flow, and one at a time, such as in CATSA chokepoint entries and other portals.
3	Unconstrained, free-flow, and many at a time, such as in airport halls, train stations, and other indoor public spaces.

Face Recognition in Video Readiness Assessment Scale

GRADE	DEFINITION
++	Operationally Ready: Deployed immediately with no customization and predictable results.
+	Operationally with Configuration: Deployed within 1 year with some customization/configuration; predictable results.
○○	Short-term: Deployed within 1 to 3 years with a moderate investment in applied research and development.
○	Medium-term: Deployed within a 3 to 5 years with a significant investment in applied research and development.
-	Academic: Deployment timeline of over 5 years; requires academic research and development.

Face Recognition in Video Readiness Evaluation

FACE RECOGNITION IN VIDEO APPLICATION	TYPE 0[1]	TYPE 1	TYPE 2	TYPE 3
Detection				
1. Face Detection in Live Video	++	++	+	○
2. Face Extraction from Archive Video	++	++	+	○
Tracking				
3. Face Tracking Across a Single Video	+	+	+	-
4. Face Tracking Across Multiple Videos	+	+	+	-
Recognition				
Still to Video				
5. Face Recognition for Watch List Screening – Binary	○○	○	-	-
6. Face Recognition for Watch List Screening – Triaging	○○	○	○	-
Video to Video				
7. Face Recognition from Multiple Videos	+	○○	○	-
8. Face Recognition Fusion from Multiple Videos	○	○	-	-
Association				
9. Assisted Face Tagging and Grouping using Visual Analytics	○○	○○	○○	○
10. Automated Face Tagging and Grouping	○○	○○	○	-
Soft Biometrics				
11. Facial Expression Analysis	+	○	○	-
12. Human attribute recognition (Gender/Age/Race)	○	○	○	-
Multiple Biometrics				
13. Face Recognition to Improve Voice/Iris Biometrics	+	○	-	-
14. Soft biometrics to improve face recognition	○	○	○	-

[1] Estimated readiness: The e-Gate environment was not evaluated in this study and is being examined separately in another project currently led by CBSA and funded under DRDC's Centre for Security Science Program.

Appendix 2: Event Detection in Video Evaluation Matrix

Event Detection in Video Environment Types

TYPE	DEFINITION
1	Controlled, such as in Primary Inspection Lanes (PIL).
2	Free-flow, and one at a time chokepoints, such as in CATSA chokepoint entries and other portals.
3	Free-flow, and many at a time chokepoints.
4	Free-flow, and many at a time in general indoor environments.
5	Free-flow, and many at a time in general outdoor environments.

Event Detection in Video Readiness Assessment Scale

GRADE	DEFINITION
++	Operationally Ready: Deployed immediately with no customization and predictable results.
+	Operationally with Configuration: Deployed within 1 year with some customization/configuration; predictable results.
○○	Short-term: Deployed within 1 to 3 years with a moderate investment in applied research and development.
○	Medium-term: Deployed within 3 to 5 years with a significant investment in applied research and development.
-	Academic: Deployment timeline of over 5 years; requires academic research and development.

Event Detection in Video Evaluation

EVENT DETECTION IN VIDEO APPLICATION	TYPE 1	TYPE 2	TYPE 3	TYPE 4	TYPE 5
1. Unattended Baggage Detection					
a. Carried Object	-	-	-	-	-
b. Dropping Object	○[1]	○[2]	-	-	-
c. Static Object (>n sec)	+	+[1]	○[2]	-	-
d. Unattended Object	○[2]	○[2]	-	-	-
e. Abandoned Object	○[2]	○[2]	-	-	-
f. Object left behind	○[2]	○[2]	-	○	-
2. Tail-gaiting Detection					
a. Tail-gating Detection	+	○	-	○	-
3. Person Tracking					
a. Person Counting	+	+	-	○	-
b. Running	+	○	-	○	-
c. Opposite directions	+	○	○	○	○
4. Person-Baggage Tagging					
a. Person Counting	○	○	-	-	-
b. Running	○	-	-	-	-
c. Opposite directions	-	-	-	-	-
5. Loitering Detection					
a. Improper standing place	+	○○	○○	○○	○
b. Wandering around	+	○	-	-	-
6. Crowd Tracking					
a. Splitting	+	+	○	-	-
b. Merging	+	+	○	-	-
c. Density estimation	○○	○○	○○	○○	○○
d. Rapid dispersion	○○	○○	○○	○○	○○
e. Crowd formation	○○	○○	○○	○○	○○
7. Camera Tampering					
a. Occlusion	++	++	++	++	++
b. Focus moved	++	++	++	++	++

[1]Good viewing conditions only
[2]Good viewing conditions, low traffic, large objects only